

#44
PCT/JP99/03722

09.07.99

REC'D 26 JUL 1999

WIPO PCT

日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

09/1743543
別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office.

出願年月日
Date of Application:

1998年 7月13日

出願番号
Application Number:

平成10年特許願第197154号

出願人
Applicant(s):

日本電気株式会社

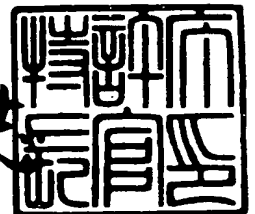
PRIORITY
DOCUMENT

SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

1999年 6月11日

特許庁長官
Commissioner,
Patent Office

伴佐山 建志



出証番号 出証特平11-3037840

【書類名】 特許願

【整理番号】 33509262

【提出日】 平成10年 7月13日

【あて先】 特許庁長官殿

【国際特許分類】 G10L 9/14

【発明の名称】 音声符号化装置ならびに音声符号化復号化装置

【請求項の数】 4

【発明者】

 【住所又は居所】 東京都港区芝五丁目7番1号 日本電気株式会社内

 【氏名】 小澤 一範

【特許出願人】

 【識別番号】 000004237

 【氏名又は名称】 日本電気株式会社

【代理人】

 【識別番号】 100082935

 【弁理士】

 【氏名又は名称】 京本 直樹

 【電話番号】 03-3454-1111

【選任した代理人】

 【識別番号】 100082924

 【弁理士】

 【氏名又は名称】 福田 修一

 【電話番号】 03-3454-1111

【選任した代理人】

 【識別番号】 100085268

 【弁理士】

 【氏名又は名称】 河合 信明

 【電話番号】 03-3454-1111

【手数料の表示】

【予納台帳番号】 008279

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9115699

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声符号化装置ならびに音声符号化復号化装置

【特許請求の範囲】

【請求項 1】

音声信号を入力しスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、過去の量子化された音源信号から適応コードブックにより遅延とゲインを求め音声信号を予測して残差を求める適応コードブック部と、前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化して出力する音源量子化部とを有する音声符号化装置において、前記音声信号から特徴を抽出してモードを判別する判別部と、前記判別部の出力があらかじめ定められたモードの場合に、音源信号を複数の非零のパルスの組合せで表わし、前記音源信号のゲインを量子化するゲインコードブックを有し、前記音声信号から前記パルスの振幅もしくは極性をあらかじめ算出し、前記パルスの位置をシフトする複数のシフト量とゲインコードブックに格納されたゲインコードベクトルとの組み合わせについて探索し、入力音声と再生信号との歪みを最小にするようにシフト量とゲインコードベクトルとの組合せを選択して出力する音源量子化部と、スペクトルパラメータ計算部の出力と判別部の出力と適応コードブック部の出力と音源量子化部の出力とを組み合わせで出力するマルチプレクサ部とを有することを特徴とする音声符号化装置。

【請求項 2】

音声信号を入力しスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、過去の量子化された音源信号から適応コードブックにより遅延とゲインを求め音声信号を予測して残差を求める適応コードブック部と、前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化して出力する音源量子化部とを有する音声符号化装置において、前記音声信号から特徴を抽出してモードを判別する判別部と、前記判別部の出力があらかじめ定められたモードの場合に音源信号を複数の非零のパルスの組合せで表わし、前記音源信号のゲインを量子化するゲインコードブックを有し、あらかじめ定められた規則により前記パルスの位置をすくなくとも 1 セット発生し、前記パルスの振幅もしくは極性を

前記音声信号からあらかじめ算出し、前記パルスの位置とゲインコードブックに格納されたゲインコードベクトルとの組み合わせについて探索し、入力音声と再生信号との歪みを最小にするように位置とゲインコードベクトルとの組合せを選択して出力する音源量子化部と、スペクトルパラメータ計算部の出力と前記判別部の出力と適応コードブック部の出力と音源量子化部の出力とを組み合わせで出力するマルチプレクサ部とを有することを特徴とする音声符号化装置。

【請求項3】

符号化側では、音声信号を入力しスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、過去の量子化された音源信号から適応コードブックにより遅延とゲインを求め音声信号を予測して残差を求める適応コードブック部と、前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化して出力する音源量子化部とを有する音声符号化装置において、前記音声信号から特徴を抽出してモードを判別する判別部と、前記判別部の出力があらかじめ定められたモードの場合に、音源信号を複数個の非零のパルスの組合せで表わし、前記音源信号のゲインを量子化するゲインコードブックを有し、前記音声信号から前記パルスの振幅もしくは極性をあらかじめ算出し、前記パルスの位置をシフトする複数のシフト量とゲインコードブックに格納されたゲインコードベクトルとの組み合わせについて探索し、入力音声と再生信号との歪みを最小にするようにシフト量とゲインコードベクトルとの組合せを選択して出力する音源量子化部と、スペクトルパラメータ計算部の出力と判別部の出力と適応コードブック部の出力と音源量子化部の出力とを組み合わせで出力するマルチプレクサ部とを有し、復号化側では、スペクトルパラメータに関する情報と判別信号に関する情報と適応コードブックに関する情報と音源信号に関する情報を入力し分離するデマルチプレクサ部と、前記判別信号があらかじめ定められたモードの場合に、音源信号を適応コードベクトルと複数個の非零のパルスの組合せと位置をシフトさせるシフト量とゲインコードベクトルから構成して発生させる音源信号発生部と、スペクトルパラメータにより構成され前記音源信号を入力し再生信号を出力する合成フィルタ部とを有することを特徴とする音声符号化復号化装置。

【請求項4】

符号化側では、音声信号を入力しスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、過去の量子化された音源信号から適応コードブックにより遅延とゲインを求め音声信号を予測して残差を求める適応コードブック部と、前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化して出力する音源量子化部とを有する音声符号化装置において、前記音声信号から特徴を抽出してモードを判別する判別部と、前記判別部の出力があらかじめ定められたモードの場合に音源信号を複数の非零のパルスの組合せで表わし、前記音源信号のゲインを量子化するゲインコードブックを有し、あらかじめ定められた規則により前記パルスの位置をすくなくとも1セット発生し、前記パルスの振幅もしくは極性を前記音声信号からあらかじめ算出し、前記パルスの位置とゲインコードブックに格納されたゲインコードベクトルとの組み合わせについて探索し、入力音声と再生信号との歪みを最小にするように位置とゲインコードベクトルとの組合せを選択して出力する音源量子化部と、スペクトルパラメータ計算部の出力と前記判別部の出力と適応コードブック部の出力と音源量子化部の出力とを組み合わせ出力するマルチプレクサ部とを有し、復号化側では、スペクトルパラメータに関する情報と判別信号に関する情報と適応コードブックに関する情報と音源信号に関する情報を入力し分離するデマルチプレクサ部と、前記判別信号があらかじめ定められたモードの場合に、音源信号を適応コードベクトルと選択されたパルスの位置に複数の非零のパルスを発生させさらにゲインコードベクトルを用いて音源信号を発生させる音源信号発生部と、スペクトルパラメータにより構成され前記音源信号を入力し再生信号を出力する合成フィルタ部とを有することを特徴とする音声符号化復号化装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、音声信号を低いビットレートで高品質に符号化するための音声符号化装置ならびに音声符号化復号化装置に関するものである。

【0002】

【従来の技術】

音声信号を高効率に符号化する方式としては、例えば、M.Schroeder and B. Atal 氏による"Code-excited linear prediction: High quality speech at very low bit rates" (Proc. ICASSP, pp.937-940, 1985 年) と題した論文 (文献 1) や、Kleijn 氏らによる"Improved speech quality and efficient vector quantization in SELP" (Proc. ICASSP, pp.155-158, 1988 年) と題した論文 (文献 2) などに記載されている CELP (Code Excited Linear Predictive Coding) が知られている。この従来例では、送信側では、フレーム毎 (例えば 20 ms) に音声信号から線形予測 (LPC) 分析を用いて、音声信号のスペクトル特性を表すスペクトルパラメータを抽出する。フレームをさらにサブフレーム (例えば 5 ms) に分割し、サブフレーム毎に過去の音源信号を基に適応コードブックにおけるパラメータ (ピッチ周期に対応する遅延パラメータとゲインパラメータ) を抽出し、適応コードブックにより前記サブフレームの音声信号をピッチ予測する。ピッチ予測して求めた音源信号に対して、予め定められた種類の雑音信号からなる音源コードブック (ベクトル量子化コードブック) から最適な音源コードベクトルを選択し、最適なゲインを計算することにより、音源信号を量子化する。音源コードベクトルの選択の仕方は、選択した雑音信号により合成した信号と、前記残差信号との誤差電力を最小化するように行う。そして、選択されたコードベクトルの種類を表すインデクスとゲインならびに、前記スペクトルパラメータと適応コードブックのパラメータをマルチプレクサ部により組み合わせて伝送する。受信側の説明は省略する。

【0003】

【発明が解決しようとする課題】

前記従来法では音源コードブックから最適な音源コードベクトルを選択するのに多大な演算量を要するという問題があった。これは、文献 1 や 2 の方法では、音源コードベクトルを選択するのに、各コードベクトルに対して一旦フィルタリングもしくは畳み込み演算を行ない、この演算をコードブックに格納されているコードベクトルの個数だけ繰り返すことに起因する。例えば、コードブックのビット数が B ビットで、次元数が N のときは、フィルタリングあるいは畳み込み演算のときのフィルタあるいはインパルス応答長を K とすると、演算量は 1 秒当

り、 $N \times K \times 2B \times 8000 / N$ だけ必要となる。一例として、 $B=10$ 、 $N=40$ 、 $k=10$ とすると、1秒当たり81,920,000回の演算が必要となり、極めて膨大であるという問題点があった。

【0004】

音源コードブック探索に必要な演算量を低減する方法として、種々のものが提案されている。例えば、ACELP (Algebraic Code Excited Linear Prediction) 方式が提案されている。これは、例えば、C.Laflamme らによる "16 kbps wideband speech coding technique based on algebraic CELP" と題した論文 (Proc. ICASSP, pp. 13-16, 1991) (文献3) 等を参照することができる。文献3の方法によれば、音源信号を複数のパルスで表し、各パルスの位置をあらかじめ定められたビット数で表し伝送する。ここで、各パルスの振幅は+1.0もしくは-1.0に限定されているため、パルス探索の演算量を大幅に低減化できる。文献3の従来法では、演算量を大幅に低減化することが可能となる。

【0005】

次に、8kb/s以上のビットレートでは良好な音質が得られるが、それ未満のビットレートでは、特に音声に背景雑音が重畳している場合に、パルスの個数が充分でなく、符号化音声の背景雑音部分の音質が極めて劣化するとい問題点があった。この理由としては、音源信号を複数のパルスの組合せで表すので、音声の母音区間では、パルスがピッチの開始点であるピッチパルスの近辺に集中するために、少ない個数のパルスで効率的に表すことができるが、背景雑音のようなランダム信号に対しては、パルスをランダムに立てる必要があるため、少ない個数のパルスでは、背景雑音を良好に表すことは困難であり、ビットレートを低減化し、パルスの個数が削減されると、背景雑音に対する音質が急激に劣化していた。

【0006】

本発明の目的は、上述の問題を解決し、ビットレートが低い場合にも、比較的少ない演算量で、特に背景雑音に対する音質の劣化の少ない音声符号化方式を提供することにある。

【0007】

【課題を解決するための手段】

本発明によれば、音声信号を入力しスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、過去の量子化された音源信号から適応コードブックにより遅延とゲインを求め音声信号を予測して残差を求める適応コードブック部と、前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化して出力する音源量子化部とを有する音声符号化装置において、前記音声信号から特徴を抽出してモードを判別する判別部と、前記判別部の出力があらかじめ定められたモードの場合に、音源信号を複数個の非零のパルスの組合せで表わし、前記音源信号のゲインを量子化するゲインコードブックを有し、前記音声信号から前記パルスの振幅もしくは極性をあらかじめ算出し、前記パルスの位置をシフトする複数のシフト量とゲインコードブックに格納されたゲインコードベクトルとの組み合わせについて探索し、入力音声と再生信号との歪みを最小にするようにシフト量とゲインコードベクトルとの組合せを選択して出力する音源量子化部と、スペクトルパラメータ計算部の出力と判別部の出力と適応コードブック部の出力と音源量子化部の出力とを組み合わせるマルチプレクサ部とを有することを特徴とする音声符号化装置が得られる。

【0008】

本発明によれば、音声信号を入力しスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、過去の量子化された音源信号から適応コードブックにより遅延とゲインを求め音声信号を予測して残差を求める適応コードブック部と、前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化して出力する音源量子化部とを有する音声符号化装置において、前記音声信号から特徴を抽出してモードを判別する判別部と、前記判別部の出力があらかじめ定められたモードの場合に音源信号を複数個の非零のパルスの組合せで表わし、前記音源信号のゲインを量子化するゲインコードブックを有し、あらかじめ定められた規則により前記パルスの位置をすくなくとも1セット発生し、前記パルスの振幅もしくは極性を前記音声信号からあらかじめ算出し、前記パルスの位置とゲインコードブックに格納されたゲインコードベクトルとの組み合わせについて探索し、入力音声と再生信号との歪みを最小にするように位置とゲインコードベクトル

ルとの組合せを選択して出力する音源量子化部と、スペクトルパラメータ計算部の出力と前記判別部の出力と適応コードブック部の出力と音源量子化部の出力とを組み合わせ出力するマルチプレクサ部とを有することを特徴とする音声符号化装置が得られる。

【0009】

本発明によれば、符号化側では、音声信号を入力しスペクトルパラメータを求めて量子化するスペクトルパラメータ計算部と、過去の量子化された音源信号から適応コードブックにより遅延とゲインを求め音声信号を予測して残差を求める適応コードブック部と、前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化して出力する音源量子化部とを有する音声符号化装置において、前記音声信号から特徴を抽出してモードを判別する判別部と、前記判別部の出力があらかじめ定められたモードの場合に、音源信号を複数の非零のパルスの組合せで表わし、前記音源信号のゲインを量子化するゲインコードブックを有し、前記音声信号から前記パルスの振幅もしくは極性をあらかじめ算出し、前記パルスの位置をシフトする複数のシフト量とゲインコードブックに格納されたゲインコードベクトルとの組み合わせについて探索し、入力音声と再生信号との歪みを最小にするようにシフト量とゲインコードベクトルとの組合せを選択して出力する音源量子化部と、スペクトルパラメータ計算部の出力と判別部の出力と適応コードブック部の出力と音源量子化部の出力とを組み合わせ出力するマルチプレクサ部とを有し、復号化側では、スペクトルパラメータに関する情報と判別信号に関する情報と適応コードブックに関する情報と音源信号に関する情報を入力し分離するデマルチプレクサ部と、前記判別信号があらかじめ定められたモードの場合に、音源信号を適応コードベクトルと複数の非零のパルスの組合せと位置をシフトさせるシフト量とゲインコードベクトルから構成して発生させる音源信号発生部と、スペクトルパラメータにより構成され前記音源信号を入力し再生信号を出力する合成フィルタ部とを有することを特徴とする音声符号化復号化装置が得られる。

【0010】

本発明によれば、符号化側では、音声信号を入力しスペクトルパラメータを求

めて量子化するスペクトルパラメータ計算部と、過去の量子化された音源信号から適応コードブックにより遅延とゲインを求め音声信号を予測して残差を求める適応コードブック部と、前記スペクトルパラメータを用いて前記音声信号の音源信号を量子化して出力する音源量子化部とを有する音声符号化装置において、前記音声信号から特徴を抽出してモードを判別する判別部と、前記判別部の出力があらかじめ定められたモードの場合に音源信号を複数の非零のパルスの組合せで表わし、前記音源信号のゲインを量子化するゲインコードブックを有し、あらかじめ定められた規則により前記パルスの位置をすくなくとも1セット発生し、前記パルスの振幅もしくは極性を前記音声信号からあらかじめ算出し、前記パルスの位置とゲインコードブックに格納されたゲインコードベクトルとの組み合わせについて探索し、入力音声と再生信号との歪みを最小にするように位置とゲインコードベクトルとの組合せを選択して出力する音源量子化部と、スペクトルパラメータ計算部の出力と前記判別部の出力と適応コードブック部の出力と音源量子化部の出力とを組み合わせ出力するマルチプレクサ部とを有し、復号化側では、スペクトルパラメータに関する情報と判別信号に関する情報と適応コードブックに関する情報と音源信号に関する情報を入力し分離するデマルチプレクサ部と、前記判別信号があらかじめ定められたモードの場合に、音源信号を適応コードベクトルと選択されたパルスの位置に複数の非零のパルスを発生させさらにゲインコードベクトルを用いて音源信号を発生させる音源信号発生部と、スペクトルパラメータにより構成され前記音源信号を入力し再生信号を出力する合成フィルタ部とを有することを特徴とする音声符号化復号化装置が得られる。

【0011】

【発明の実施の形態】

図1は本発明による音声符号化装置の一実施例を示すブロック図である。

【0012】

図において、入力端子100から音声信号を入力し、フレーム分割回路110では音声信号をフレーム（例えば20ms）毎に分割し、サブフレーム分割回路120では、フレームの音声信号をフレームよりも短いサブフレーム（例えば5ms）に分割する。

【0013】

スペクトルパラメータ計算回路200では、少なくとも一つのサブフレームの音声信号に対して、サブフレーム長よりも長い窓（例えば24ms）をかけて音声を取り出してスペクトルパラメータをあらかじめ定められた次数（例えばP=10次）計算する。ここでスペクトルパラメータの計算には、周知のLPC分析や、Burg分析等を用いることができる。ここでは、Burg分析を用いることとする。Burg分析の詳細については、中溝著による“信号解析とシステム同定”と題した単行本（コロナ社1988年刊）の82～87頁（文献4）等に記載されているので説明は略する。さらにスペクトルパラメータ計算部では、Burg法により計算された線形予測係数 α_i （ $i=1, \dots, 10$ ）を量子化や補間に適したLSPパラメータに変換する。ここで、線形予測係数からLSPへの変換は、菅村他による“線スペクトル対（LSP）音声分析合成方式による音声情報圧縮”と題した論文（電子通信学会論文誌、J64-A, pp.599-606, 1981年）（文献5）を参照することができる。例えば、第2、4サブフレームでBurg法により求めた線形予測係数を、LSPパラメータに変換し、第1、3サブフレームのLSPを直線補間により求めて、第1、3サブフレームのLSPを逆変換して線形予測係数に戻し、第1-4サブフレームの線形予測係数 α_{i1} （ $i=1, \dots, 10$ 、 $l=1, \dots, 5$ ）を聴感重み付け回路230に出力する。また、第4サブフレームのLSPをスペクトルパラメータ量子化回路210へ出力する。

【0014】

スペクトルパラメータ量子化回路210では、あらかじめ定められたサブフレームのLSPパラメータを効率的に量子化し、下式の歪みを最小化する量子化値を出力する。

【0015】

【数1】

$$D_j = \sum_{i=1}^{10} W(i) [LSP(i) - QLSP(i)_j]^2 \quad (1)$$

【0016】

ここで、 $LSP(i)$ 、 $QLSP(i)j$ 、 $W(i)$ はそれぞれ、量子化前の*i*次目のLSP、量子化後の*j*番目の結果、重み係数である。

【0017】

以下では、量子化法として、ベクトル量子化を用いるものとし、第4サブフレームのLSPパラメータを量子化するものとする。LSPパラメータのベクトル量子化の手法は周知の手法を用いることができる。具体的な方法は例えば、特開平4-171500号公報（特願平2-297600号）（文献6）や特開平4-363000号公報（特願平3-261925号）（文献7）や、特開平5-6199号公報（特願平3-155049号）（文献8）や、T.Nomura et al., による”LSP Coding Using VQ-SVQ With Interpolation in 4.075 kbps M-LCELP Speech Coder”と題した論文(Proc. Mobile Multimedia Communications, pp.B.2.5,1993)（文献9）等を参照できるのでここでは説明は略する。

【0018】

また、スペクトルパラメータ量子化回路210では、第4サブフレームで量子化したLSPパラメータをもとに、第1～第4サブフレームのLSPパラメータを復元する。ここでは、現フレームの第4サブフレームの量子化LSPパラメータと1つ過去のフレームの第4サブフレームの量子化LSPを直線補間して、第1～第3サブフレームのLSPを復元する。ここで、量子化前のLSPと量子化後のLSPとの誤差電力を最小化するコードベクトルを1種類選択した後に、直線補間により第1～第4サブフレームのLSPを復元できる。さらに性能を向上させるためには、前記誤差電力を最小化するコードベクトルを複数候補選択したのちに、各々の候補について、累積歪を評価し、累積歪を最小化する候補と補間LSPの組を選択するようにすることができる。詳細は、例えば、特願平5-8737号明細書（文献10）を参照することができる。

【0019】

以上により復元した第1-3サブフレームのLSPと第4サブフレームの量子化LSPをサブフレーム毎に線形予測係数 α_{il} （ $i=1, \dots, 10$, $l=1, \dots, 5$ ）に変換し、インパルス応答計算回路310へ出力する。また、第4サブ

フレームの量子化LSPのコードベクトルを表すインデックスをマルチプレクサ400に出力する。

【0020】

聴感重み付け回路230は、スペクトルパラメータ計算回路200から、各サブフレーム毎に量子化前の線形予測係数 α_{i1} ($i=1, \dots, 10, l=1, \dots, 5$)を入力し、前記文献1にもとづき、サブフレームの音声信号に対して聴感重み付けを行い、聴感重み付け信号を出力する。

【0021】

応答信号計算回路240は、スペクトルパラメータ計算回路200から、各サブフレーム毎に線形予測係数 α_{i1} を入力し、スペクトルパラメータ量子化回路210から、量子化、補間して復元した線形予測係数 α_{i1} をサブフレーム毎に入力し、保存されているフィルタメモリの値を用いて、入力信号を零 $d(n)=0$ とした応答信号を1サブフレーム分計算し、減算器235へ出力する。ここで、応答信号 $x_z(n)$ は下式で表される。

【0022】

【数2】

$$x_z(n) = d(n) - \sum_{i=1}^{10} \alpha_i d(n-i) + \sum_{i=1}^{10} \alpha_i \gamma^i y(n-i) + \sum_{i=1}^{10} \alpha_i \gamma^i x_z(n-i) \quad (2)$$

ただし、 $n-i \leq 0$ のときは

$$y(n-i) = p(N+(n-i)) \quad (3)$$

$$x_z(n-i) = s_w(N+(n-i)) \quad (4)$$

【0023】

ここでNはサブフレーム長を示す。 γ は、聴感重み付け量を制御する重み係数であり、下記の式(7)と同一の値である。 $s_w(n)$ 、 $p(n)$ は、それぞれ

、重み付け信号計算回路の出力信号、後述の式(7)における右辺第1項のフィルタの分母の項の出力信号をそれぞれ示す。

【0024】

減算器235は、下式により、聴感重み付け信号から応答信号を1サブフレーム分減算し、 $x'_w(n)$ を適応コードブック回路300へ出力する。

【0025】

【数3】

$$x'_w(n) = x_w(n) - x_z(n) \quad (5)$$

【0026】

インパルス応答計算回路310は、 z 変換が下式で表される聴感重み付けフィルタのインパルス応答 $h_w(n)$ をあらかじめ定められた点数 L だけ計算し、適応コードブック回路500、音源量子化回路350へ出力する。

【0027】

【数4】

$$H_w(z) = \frac{1 - \sum_{i=1}^{10} \alpha_i z^{-i}}{1 - \sum_{i=1}^{10} \alpha_i \gamma^i z^{-i}} \frac{1}{1 - \sum_{i=1}^{10} \alpha_i \gamma^i z^{-i}} \quad (6)$$

【0028】

モード判別回路800は、フレーム分割回路の出力信号を用いて、特徴量を抽出し、フレーム毎にモードの判別を行う。ここで、特徴としては、ピッチ予測ゲインを用いることができる。サブフレーム毎に求めたピッチ予測ゲインをフレーム全体で平均し、この値とあらかじめ定められた複数のしきい値を比較し、あらかじめ定められた複数のモードに分類する。ここでは、一例として、モードの種類は4とする。この場合、モード0、1、2、3は、それぞれ、無声区間、過渡区間、弱い有声区間、強い有声区間にほぼ対応するものとする。モード判別情報を音源量子化回路350とゲイン量子化回路365とマルチプレクサ400へ出

力する。

【0029】

適応コードブック回路500では、ゲイン量子化回路365から過去の音源信号 $v(n)$ を、減算器235から出力信号 $x'_w(n)$ を、インパルス応答計算回路310から聴感重み付けインパルス応答 $h_w(n)$ を入力する。ピッチに対応する遅延 T を下式の歪みを最小化するように求め、遅延を表すインデックスをマルチプレクサ400に出力する。

【0030】

【数5】

$$D_T = \sum_{n=0}^{N-1} x_w'^2(n) - \left[\sum_{n=0}^{N-1} x_w'(n) y_w(n-T) \right]^2 / \left[\sum_{n=0}^{N-1} y_w^2(n-T) \right] \quad (7)$$

ただし

$$y_w(n-T) = v(n-T) * h_w(n) \quad (8)$$

【0031】

式(8)において、記号*は畳み込み演算を表す。

【0032】

次に、ゲイン β を下式に従い求める。

【0033】

【数6】

$$\beta = \sum_{n=0}^{N-1} x_w'(n) y_w(n-T) / \sum_{n=0}^{N-1} y_w^2(n-T) \quad (9)$$

【0034】

ここで、女性音や、子供の声に対して、遅延の抽出精度を向上させるために、遅延を整数サンプルではなく、小数サンプル値で求めてもよい。具体的な方法は、例えば、P.Kroon らによる、"Pitch pre-dictors with high temporal resolu

tion”と題した論文 (Proc. ICASSP, pp.661-664, 1990 年) (文献 11) 等を参照することができる。さらに、適応コードブック回路 500 では式 (10) に従いピッチ予測を行ない、予測残差信号 $e_w(n)$ を音源量子化回路 350 へ出力する。

【0035】

【数 7】

$$e_w(n) = x_w(n) - \beta v(n-T) * h_w(n) \quad (10)$$

【0036】

音源量子化回路 350 では、モード判別情報を入力し、モードにより、音源信号の量子化方法を切り替える。

【0037】

モード 1、2、3 では、M 個のパルスをとてるとものする。モード 1、2、3 では、パルスの振幅を M パルス分まとめて量子化するための、B ビットの振幅コードブック、もしくは極性コードブックを有しているものとする。以下では、極性コードブックを用いる場合の説明を行なう。この極性コードブックは、音源コードブック 351 に格納されている。

【0038】

有声では、音源量子化回路 350 は、コードブック 351 に格納された各極性コードベクトルを読みだし、各コードベクトルに対して位置をあてはめ、式 (11) を最小化するコードベクトルと位置の組合せを複数セット選択する。

【0039】

【数 8】

$$D_k = \sum_{n=0}^{N-1} [e_w(n) - \sum_{i=1}^M g_{ik} h_w(n-m_i)]^2 \quad (11)$$

【0040】

ここで、 $h_w(n)$ は、聴感重み付けインパルス応答である。

式(11)を最小化するには、式(12)を最大化する極性コードベクトル g_i と位置 m_i の組合せを求めれば良い。

【0041】

【数9】

$$D_{(k,i)} = \left[\sum_{n=0}^{N-1} e_w(n) s_{wk}(m_i) \right]^2 / \sum_{n=0}^{N-1} s_{wk}^2(m_i) \quad (12)$$

【0042】

または、式(13)を最大化するように選択しても良い。この方が分子の計算に要する演算量が低減化される。

【0043】

【数10】

$$D_{(k,i)} = \left[\sum_{n=0}^{N-1} \Phi(n) v_k(n) \right]^2 / \sum_{n=0}^{N-1} s_{wk}^2(m) \quad (13)$$

$$\text{ただし } \Phi(n) = \sum_{i=n}^{N-1} e_w(i) h_w(i-n), n=0, \dots, N-1 \quad (14)$$

【0044】

ここで、モード1-3の場合の各パルスのとおり得る位置は、演算量削減のため、文献3に示すように、拘束することができる。一例として、 $N=40$ 、 $M=5$ とすると、各パルスのとおり得る位置は表1のようになる。

【0045】

【表 1】

パルス番号	位置
第 1 パルス	0, 5, 10, 15, 20, 25, 30, 35
第 2 パルス	1, 6, 11, 16, 21, 26, 31, 36
第 3 パルス	2, 7, 12, 17, 22, 27, 32, 37
第 4 パルス	3, 8, 13, 18, 23, 28, 33, 38
第 5 パルス	4, 9, 14, 19, 24, 29, 34, 39

【0046】

極性コードベクトルの探索終了後、選択された複数セットの極性コードベクトルと位置の組み合わせをゲイン量子化回路 365 に出力する。

【0047】

あらかじめ定められたモード（この例ではモード 0）では、表 2 のように、パルスの位置を一定の間隔で定め、パルス全体の位置をシフトさせるための複数のシフト量をさだめておく。以下の例の場合は、位置を 1 サンプルずつシフトさせるとして、4 種類のシフト量（シフト 0, シフト 1, シフト 2, シフト 3）を用いる。また、この場合はシフト量を 2 ビットで量子化して伝送する。表 2 において、シフト量 0 の場合は基本的なパルスの位置を表す。シフト量 1、2、3 の場合は、シフト量 0 の場合のパルス位置を一律にそれぞれ、1 サンプル、2 サンプル、3 サンプルシフトしたものである。これらの 4 種類のシフト量を本実施例では用いることにするが、シフト量の種類、シフトサンプル数は任意に設定できる。

【0048】

【表 2】

シフト量	位置
0	0, 4, 8, 12, 16, 20, 24, 28, 32, 36
1	1, 5, 9, 13, 17, 21, 25, 29, 33, 37
2	2, 6, 10, 14, 18, 22, 26, 30, 34, 38
3	3, 7, 11, 15, 19, 23, 27, 31, 35, 39

【0049】

表 2 の各シフト量及び各パルス位置に対する極性を、式 (14) からあらかじめ求めておく。

【0050】

各シフト量毎に、表 2 に示す位置とそれに対応する極性を、ゲイン量子化回路 365 に出力する。

【0051】

ゲイン量子化回路 365 は、モード判別回路 800 からモード判別情報を入力する。音源量子化回路 350 から、モード 1-3 では、複数セットの極性コードベクトルとパルス位置の組み合わせを入力し、モード 0 では、シフト量毎にパルスの位置とそれに対応する極性の組み合わせを入力する。

【0052】

ゲイン量子化回路 365 は、ゲインコードブック 380 からゲインコードベクトルを読みだし、モード 1-3 では、選択された複数セットの極性コードベクトルと位置の組み合わせに対して、式 (15) を最小化するようにゲインコードベクトルを探索し、歪みを最小化するゲインコードベクトル、極性コードベクトルと位置の組み合わせを 1 種類選択する。

【0053】

【数11】

$$D_k = \sum_{n=0}^{N-1} [x_w(n) - \beta_k v(n-T) * h_w(n) - G'_k \sum_{i=1}^M g_{ik} h_w(n-m_i)]^2 \quad (15)$$

【0054】

ここでは、適応コードブックのゲインとパルスで表した音源のゲインの両者を同時にベクトル量子化する例について示した。選択された極性コードベクトルを表すインデクス、位置を表す符号、ゲインコードベクトルを表すインデクスをマルチプレクサ400に出力する。

【0055】

判別情報がモード0の場合は、複数のシフト量と各シフト量の場合の各位置に対応した極性を入力し、ゲインコードベクトルを探索し、式(16)を最小化するようにゲインコードベクトルとシフト量を1種類選択する。

【0056】

【数12】

$$D_k = \sum_{n=0}^{N-1} [x_w(n) - \beta_k v(n-T) * h_w(n) - G'_k \sum_{i=1}^M g_{ik} h_w(n-m_i - \delta(j))]^2 \quad (16)$$

【0057】

ここで、 β_k 、 G'_k は、ゲインコードブック380に格納された2次元ゲインコードブックにおけるk番目のコードベクトルである。また、 $\delta(j)$ はj番目のシフト量を示し、 g'_k は選択されたゲインコードベクトルを表す。選択されたゲインコードベクトルを表すインデクスとシフト量を表す符号をマルチプレクサ400に出力する。

【0058】

なお、モード1-3では、複数パルスの振幅を量子化するためのコードブックを、音声信号を用いてあらかじめ学習して格納しておくこともできる。コードブ

ックの学習法は、例えば、Linde 氏らによる "An algorithm for vector quantization design," と題した論文 (IEEE Trans. Commun., pp.84-95, January, 1980) (文献 12) 等を参照できる。

【0059】

重み付け信号計算回路 360 は、モード判別情報とそれぞれのインデックスを入力し、インデックスからそれに対応するコードベクトルを読みだす。モード 1-3 の場合は、式 (17) にもとづき駆動音源信号 $v(n)$ を求める。

【0060】

【数 13】

$$v(n) = \beta'_{\tau} v(n-T) + G'_{\tau} \sum_{i=1}^M g'_{ik} \delta(n-m_i) \quad (17)$$

【0061】

$v(n)$ は適応コードブック回路 500 に出力される。

【0062】

モード 1-3 の場合は、式 (18) にもとづき駆動音源信号 $v(n)$ を求める。

【0063】

【数 14】

$$v(n) = \beta'_{\tau} v(n-T) + G'_{\tau} \sum_{i=1}^M g'_{ik} \delta(n-m_i - \delta(j)) \quad (18)$$

【0064】

$v(n)$ は適応コードブック回路 500 に出力される。

【0065】

次に、スペクトルパラメータ計算回路 200 の出力パラメータ、スペクトルパ

ラメータ量子化回路 210 の出力パラメータを用いて式 (19) により、応答信号 $s_w(n)$ をサブフレーム毎に計算し、応答信号計算回路 240 へ出力する。

【0066】

【数 15】

$s_w(n)$

$$= v(n) - \sum_{i=1}^{10} \alpha_i v(n-i) + \sum_{i=1}^{10} \alpha_i \gamma^i p(n-i) + \sum_{i=1}^{10} \alpha_i \gamma^i s_w(n-i) \quad (19)$$

【0067】

以上により、第 1 の発明に対応する実施例の説明を終える。

【0068】

第 2 の実施例を示すブロック図を図 2 に示す。図 2 において、図 1 と同一の番号を付した構成要素は、図 1 と同一の動作を行うので、説明を省略する。

【0069】

図 2 においては、音源量子化回路 355 の動作が異なる。ここでは、モード判別情報がモード 0 の場合に、パルスの位置として、あらかじめ定められた規則に従い発生した位置を使用する。

【0070】

例えば、あらかじめ定められた個数（例えば M1）のパルスの位置を乱数発生回路 600 により発生させる。つまり、乱数発生器により発生された M1 個の数値をパルスの位置と考える。さらにこの位置のセットを複数種類発生させる。これにより発生された複数セット分の M1 個の位置を音源量子化回路 355 に出力する。

【0071】

音源量子化回路 355 は、モード判別情報がモード 1-3 の場合は、図 1 の音源量子化回路 350 と同一の動作を行なう。モード 0 の場合は、乱数発生回路 600 から出力された複数セットの位置の各々に対して、式 (14) から極性をあ

らかじめ計算する。

【0072】

複数セットの位置と各々のパルス位置に対応する極性を、ゲイン量子化回路 370 へ出力する。

【0073】

ゲイン量子化回路 370 は、複数セットの位置と各々のパルス位置に対応する極性を入力し、ゲインコードブック 380 に格納されたゲインコードベクトルを組み合わせ探索し、式 (20) を最小化するような位置のセットとゲインコードベクトルの組み合わせを 1 種類選択して出力する。

【0074】

【数 16】

$D_{k,j}$

$$= \sum_{n=0}^{N-1} [x_w(n) - \beta_i(n-T) * h_w(n) - G_i \sum_{i=1}^M g_{ik} h_w(n - m_i - \delta(j))]^2 \quad (20)$$

【0075】

以上で第 2 の発明の説明を終了する。

【0076】

図 3、図 4 は第 3 の実施例を示すブロック図である。図 3 は符号化側を示し、図 4 は復号化側を示す。図 3、図 4 において、図 1 と同一の番号を付した構成要素は、図 1 と同一の動作を行うので、説明は省略する。

【0077】

図 4 において、デマルチプレクサ 500 は、受信した信号から、モード判別情報、ゲインコードベクトルを示すインデックス、適応コードブックの遅延を示すインデックス、音源信号の情報、音源コードベクトルのインデックス、スペクトルパラメータのインデックスを入力し、各パラメータを分離して出力する。

【0078】

ゲイン復号回路 510 は、ゲインコードベクトルのインデクスとモード判別情報を入力し、ゲインコードブック 380 からインデクスに応じてゲインコードベクトルを読み出し、出力する。

【0079】

適応コードブック回路 520 は、モード判別情報と適応コードブックの遅延を入力し、適応コードベクトルを発生し、ゲインコードベクトルにより適応コードブックのゲインを乗じて出力する。

【0080】

音源信号復元回路 540 では、モード判別情報がモード 1-3 のときは、音源コードブック 351 から読み出した極性コードベクトルと、パルスの位置情報とゲインコードベクトルを用いて、音源信号を発生して加算器 550 に出力する。モード判別情報がモード 0 の場合は、パルス位置、位置のシフト量とゲインコードベクトルから音源信号を発生して加算器 550 に出力する。

【0081】

加算器 550 は、適応コードブック回路 520 の出力と音源信号復元回路 540 の出力を用いて、モード 1-3 の場合は式 (17) にもとづき、モード 0 の場合は式 (18) にもとづき駆動音源信号 $v(n)$ を発生し、適応コードブック回路 520 と合成フィルタ 560 に出力する。

【0082】

スペクトルパラメータ復号回路 570 は、スペクトルパラメータを復号し、線形予測係数に変換し、合成フィルタ回路 560 に出力する。

【0083】

合成フィルタ回路 560 は、駆動音源信号 $v(n)$ と線形予測係数を入力し、再生信号を計算し端子 580 から出力する。

【0084】

以上で第 3 の実施例の説明を終える。

【0085】

図 5、図 6 は第 4 の実施例を示すブロック図である。図 5 は符号化側を示し、図 6 は復号化側を示す。図 5、図 6 において、図 2、図 3、図 4 と同一の番号を

付した構成要素は、同一の動作をするので、説明は省略する。

【0086】

図において、音源信号復元回路 590 は、モード判別情報がモード 1-3 のときは、音源コードブック 351 から読み出した極性コードベクトルと、パルスの位置情報とゲインコードベクトルを用いて、音源信号を発生して加算器 550 に出力する。モード判別情報がモード 0 の場合は、乱数発生器 600 からパルスの位置を発生させ、ゲインコードベクトルを用いて音源信号を発生して加算器 550 に出力する。

【0087】

以上で第 4 の実施例の説明を終える。

【0088】

【発明の効果】

以上説明したように、本発明によれば、音声信号から特徴量をもとに、モードを判別し、あらかじめ定められたモードの場合に、非零の振幅のパルスにより音源信号を表し、パルス位置の振幅もしくは極性を入力音声信号からあらかじめ計算し、複数種のシフト量とゲインコードベクトルとの組み合わせを探索し、再生信号と入力音声との歪みを最小にするゲインコードベクトルとシフト量の組合せを 1 種類選択している。

【0089】

また、本発明によれば、あらかじめ定められたモードの場合に、非零の振幅のパルスにより音源信号を表し、あらかじめ定められた規則により発生した複数セットの位置に対応する振幅もしくは極性を入力音声信号から計算し、前記複数セットの位置とゲインを量子化するためのゲインコードブックに格納されるゲインコードベクトルとを組み合わせ探索し、再生信号と入力音声との歪みを最小にするゲインコードベクトルと位置のセットとの組合せを選択している。

【0090】

これらの構成により、あらかじめ定められたモードにおいて、従来方式に比べパルスの個数を大幅に増やすことができるので、背景雑音が重畳した音声を低ビットレートで符号化しても、背景雑音部分が良好に符号化できるという効果があ

る。

【図面の簡単な説明】

【図 1】

第 1 の実施例を示す図

【図 2】

第 2 の実施例を示す図

【図 3】

第 3 の実施例を示す図

【図 4】

第 3 の実施例を示す図

【図 5】

第 4 の実施例を示す図

【図 6】

第 4 の実施例を示す図

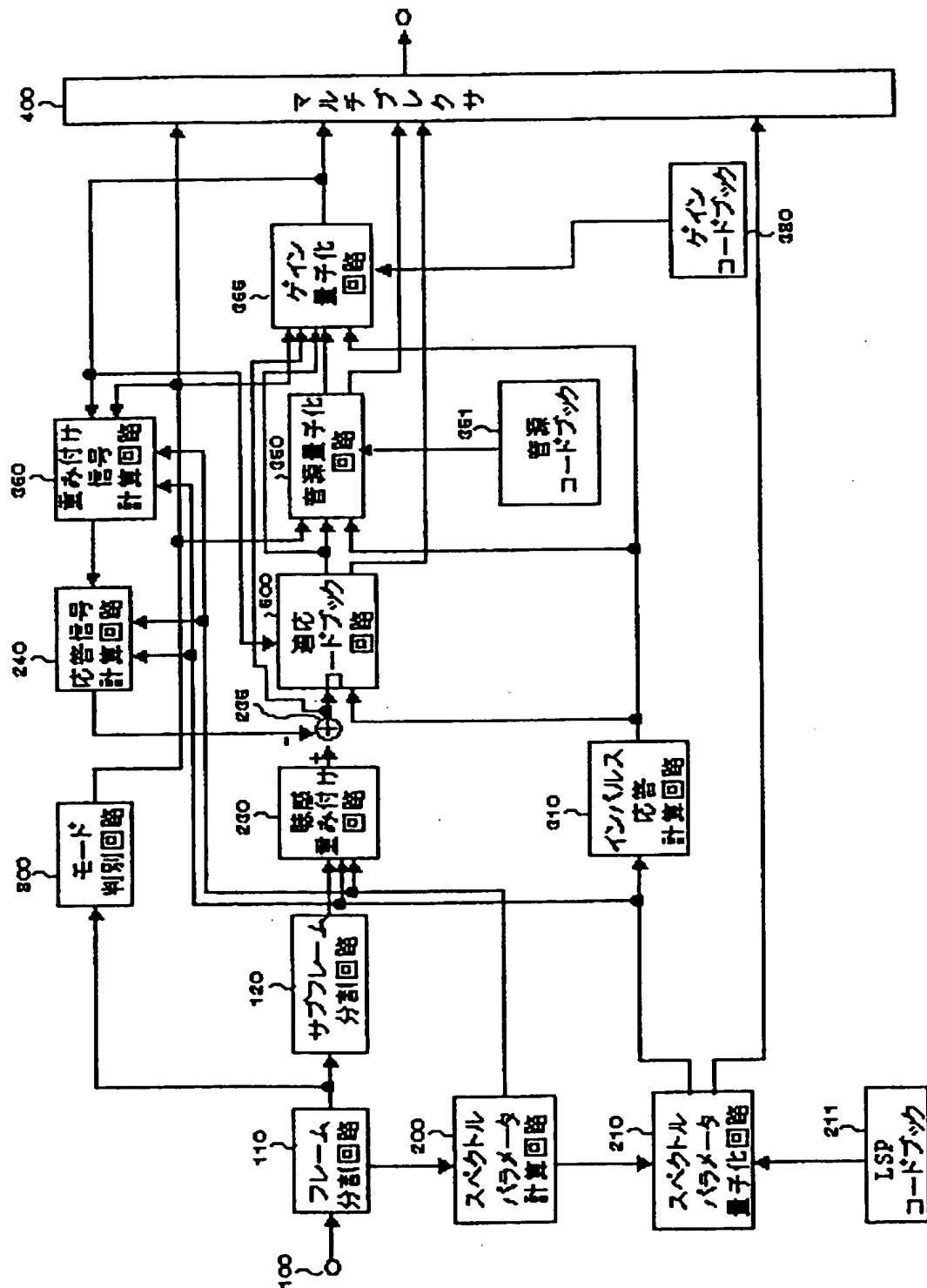
【符号の説明】

- 110 フレーム分割回路
- 120 サブフレーム分割回路
- 200 スペクトルパラメータ計算回路
- 210 スペクトルパラメータ量子化回路
- 211 LSPコードブック
- 230 聴感重み付け回路
- 235 減算回路
- 240 応答信号計算回路
- 310 インパルス応答計算回路
- 350、355、356、357 音源量子化回路
- 351 音源コードブック
- 360 重み付け信号計算回路
- 365、370 ゲイン量子化回路
- 380 ゲインコードブック

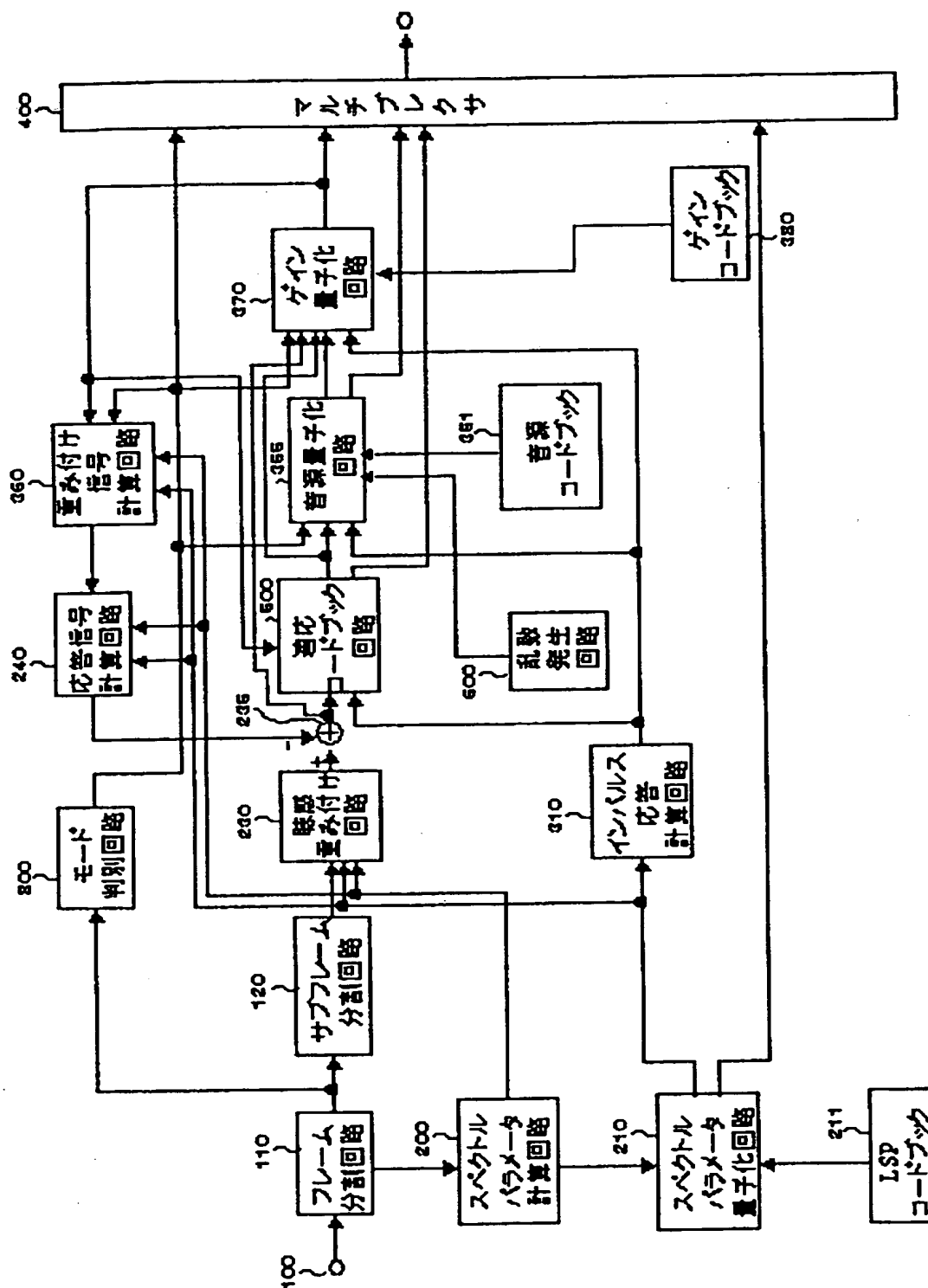
- 400 マルチプレクサ
- 500 適応コードブック回路
- 510 デマルチプレクサ
- 510 ゲイン復号回路
- 520 適応コードブック回路
- 540 音源信号復元回路
- 550 加算回路
- 560 合成フィルタ回路
- 570 スペクトルパラメータ復号回路
- 600 乱数発生回路
- 800 モード判別回路

【書類名】 図面

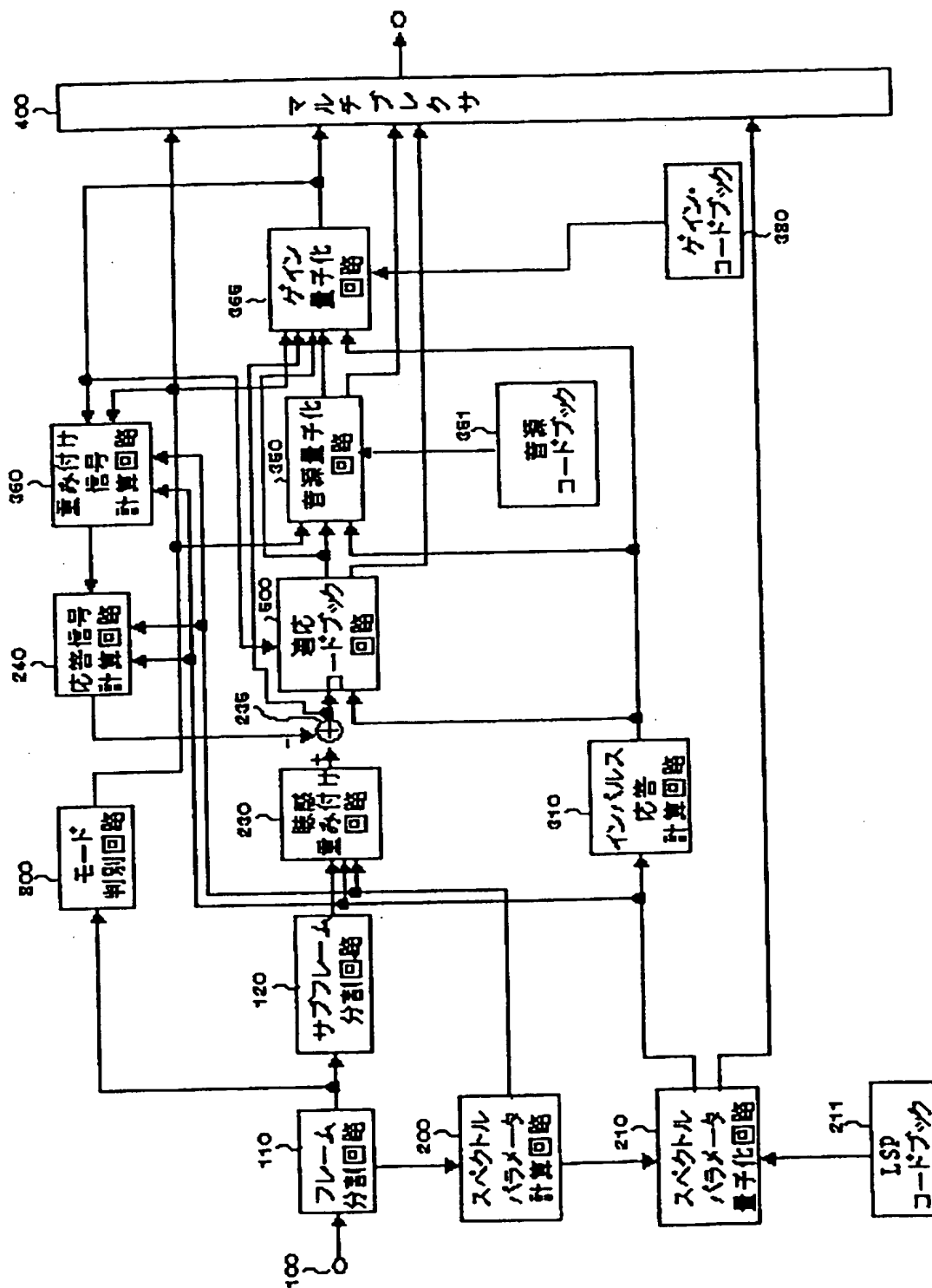
【図 1】



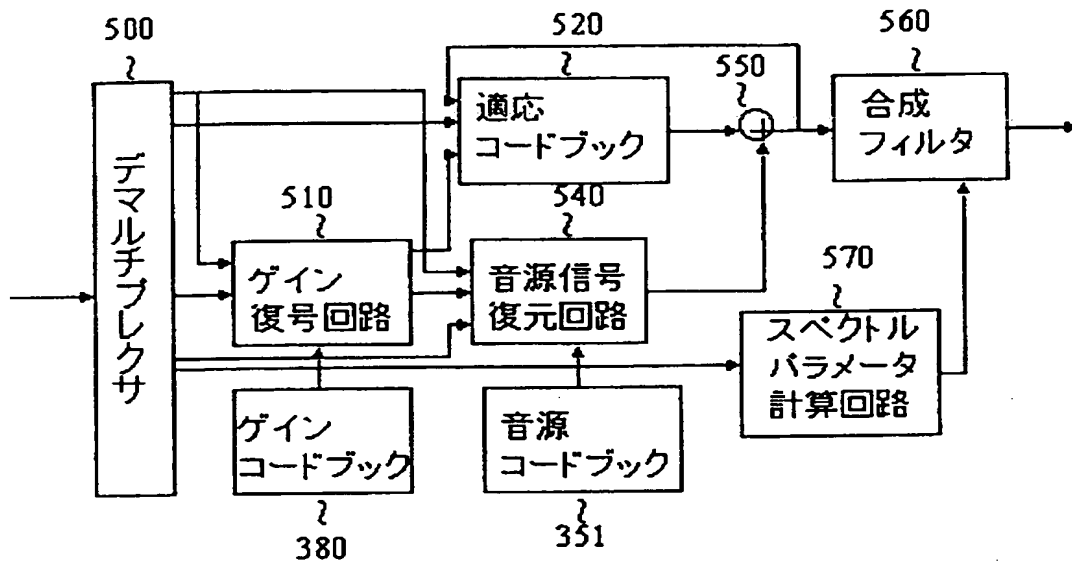
【図2】



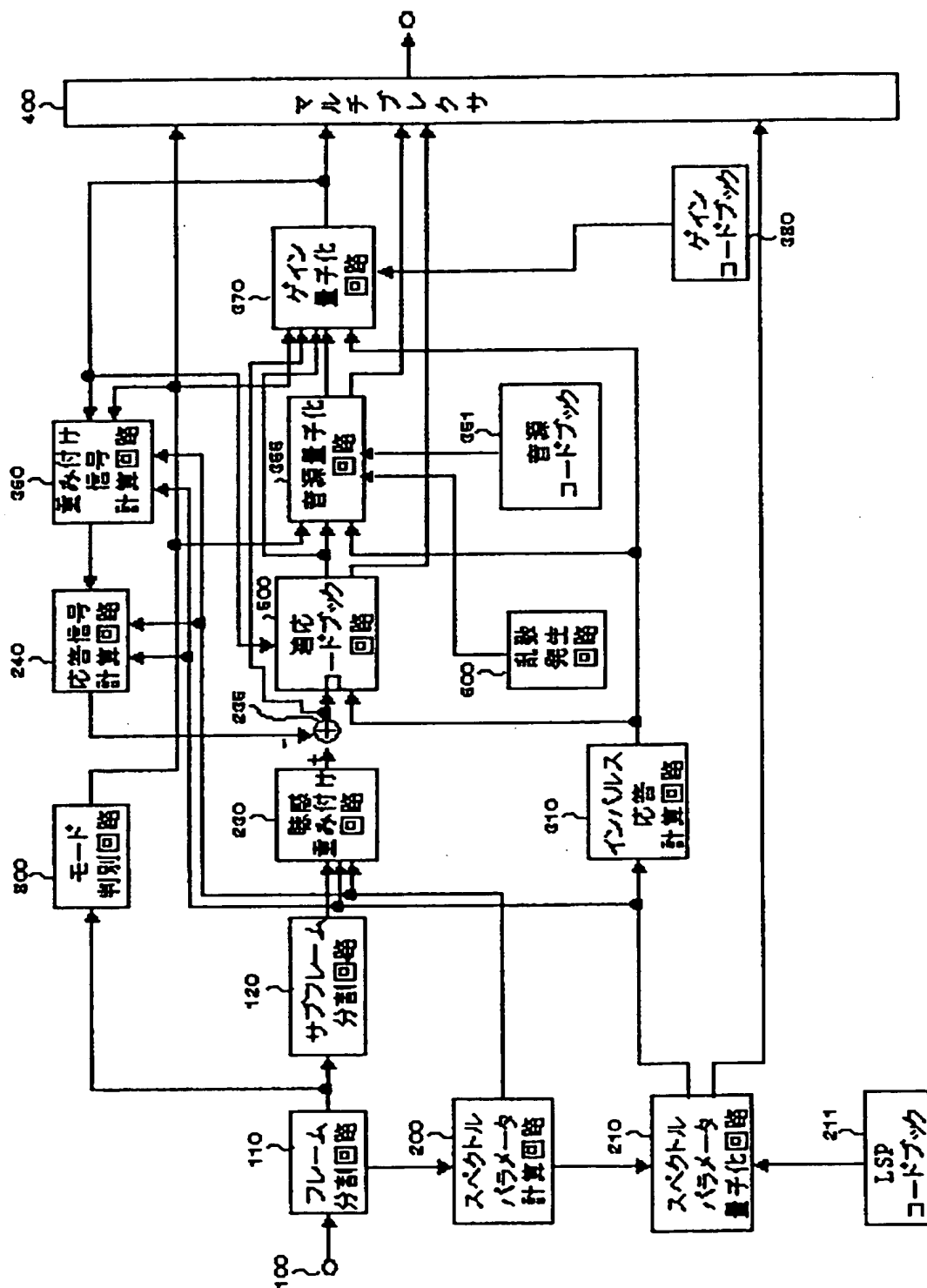
【図 3】



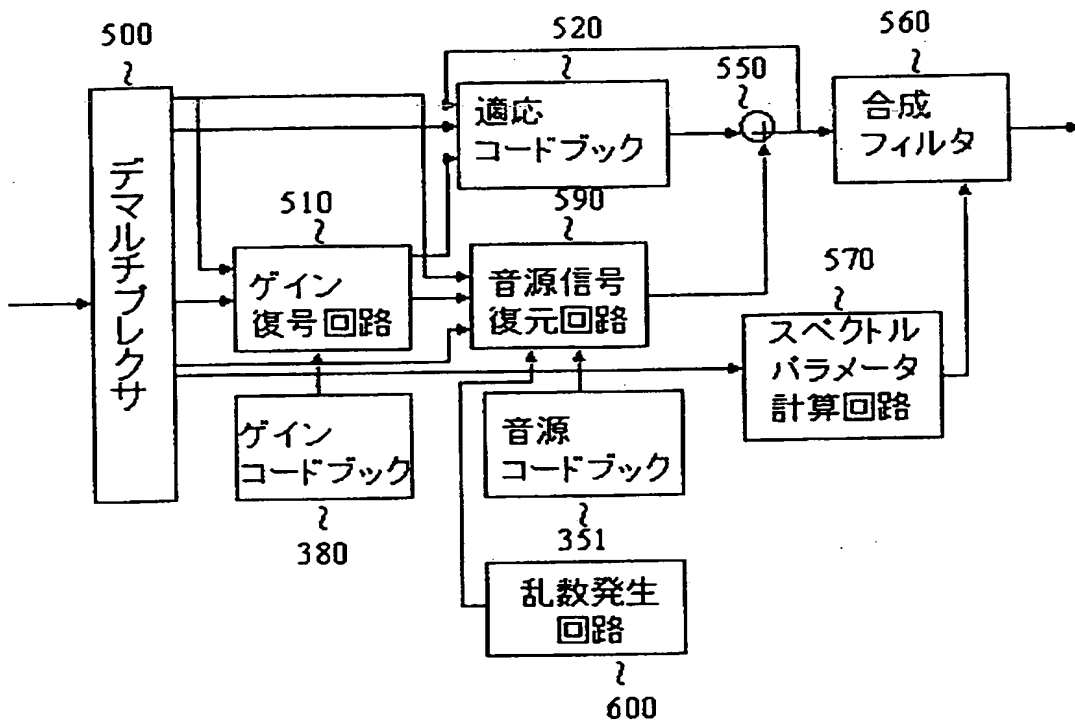
【図4】



【図 5】



【図 6】



【書類名】 要約書

【要約】

【課題】 低ビットレートでも良好な音質の得られる音声符号化装置の提供。

【解決手段】 音声符号化装置のモード判別回路 800 において、サブフレーム毎に入力音声信号から特徴量を用いてモードを判別し、音源量子化回路 350 において、あらかじめ定められたモードの場合に、非零のパルスの振幅もしくは極性をあらかじめ計算し、あらかじめ定められたパルスの位置を時間的にシフトする複数のシフト量とゲインを量子化するゲインコードベクトルとの組み合わせを探索し、再生音声と入力音声との歪みを最小にするゲインコードベクトルとシフト量の組合せを選択する。

【選択図】 図 1

【書類名】 職権訂正データ
【訂正書類】 特許願

<認定情報・付加情報>

【特許出願人】

【識別番号】 000004237
【住所又は居所】 東京都港区芝五丁目7番1号
【氏名又は名称】 日本電気株式会社

【代理人】

申請人

【識別番号】 100082935
【住所又は居所】 東京都港区芝5丁目7番1号 日本電気株式会社内
【氏名又は名称】 京本 直樹

【選任した代理人】

【識別番号】 100082924
【住所又は居所】 東京都港区芝5丁目7番1号 日本電気株式会社内
【氏名又は名称】 福田 修一

【選任した代理人】

【識別番号】 100085268
【住所又は居所】 東京都港区芝5丁目7番1号 日本電気株式会社内
【氏名又は名称】 河合 信明

出 願 人 履 歴 情 報

識別番号 [000004237]

1. 変更年月日	1990年 8月29日
[変更理由]	新規登録
住 所	東京都港区芝五丁目7番1号
氏 名	日本電気株式会社

This Page Blank (uspio)